

SEQUENCE LISTING

(1) GENERAL INFORMATION:



(i) APPLICANT: TAKEMURA, FUMINORI  
UENO, EIICHI  
ITOH, SATORU

(ii) TITLE OF INVENTION: NUCLEIC ACID-BOUND POLYPEPTIDE, METHOD  
OF PRODUCING NUCLEIC ACID-BOUND POLYPEPTIDE AND  
IMMUNOASSAY USING THE POLYPEPTIDE.

(iii) NUMBER OF SEQUENCES: 20

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
P.C.  
(B) STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
(C) CITY: ARLINGTON  
(D) STATE: VA  
(E) COUNTRY: U.S.A.  
(F) ZIP: 22202

a' (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/841,657  
(B) FILING DATE: 30-APR-1997  
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: JP 8-134444  
(B) FILING DATE: 01-MAY-1997

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: OBLON, NORMAN F.  
(B) REGISTRATION NUMBER: 24,618  
(C) REFERENCE/DOCKET NUMBER: 2084-033-0

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (703) 413-3000  
(B) TELEFAX: (703) 413-2220

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 102 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..102

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AGA	CGA	CGA	GGG	AGG	TCC	CCT	AGA	AGA	AGA	ACT	CCC	TCG	CCT	CGC	AGA	48
Arg	Arg	Arg	Gly	Arg	Ser	Pro	Arg	Arg	Arg	Thr	Pro	Ser	Pro	Arg	Arg	
1				5					10					15		
CGA	AGG	TCT	AAA	TCG	CCG	CGT	CGC	AGA	AGA	TCT	CAA	TCT	CGG	GAA	TCT	96
Arg	Arg	Ser	Lys	Ser	Pro	Arg	Arg	Arg	Arg	Ser	Gln	Ser	Arg	Glu	Ser	
			20					25					30			
CAA	TGT															102
Gln	Cys															

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Arg	Arg	Arg	Gly	Arg	Ser	Pro	Arg	Arg	Arg	Thr	Pro	Ser	Pro	Arg	Arg
1				5					10					15	
Arg	Arg	Ser	Lys	Ser	Pro	Arg	Arg	Arg	Arg	Ser	Gln	Ser	Arg	Glu	Ser
			20					25					30		

Gln Cys

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 360 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..360

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATG AGC ACA AAT CCT AAA CCT CAA AGA AAA ACC AAA CGT AAC ACC AAC	48
Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn	
1 5 10 15	
CGC CGC CCA CAG GAC GTT AAG TTC CCG GGC GGT GGT CAG ATC GTT GGT	96
Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly	
20 25 30	
GGA GTT TAC CTG TTG CCG CGC AGG GGC CCC AGG TTG GGT GTG CGC GCG	144
Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala	
35 40 45	
ACT AGG AAG ACT TCC GAG CGG TCG CAA CCT CGT GGA AGG CGA CAA CCT	192
Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro	
50 55 60	
ATC CCC AAG GCT CGC CGG CCC GAG GGT AGG ACC TGG GCT CAG CCC GGG	240
Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Thr Trp Ala Gln Pro Gly	
65 70 75 80	
TAC CCT TGG CCC CTC TAT GGC AAC GAG GGT ATG GGG TGG GCA GGA TGG	288
Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Met Gly Trp Ala Gly Trp	
85 90 95	
CTC CTG TCA CCC CGT GGC TCT CGG CCT AGT TGG GGC CCC ACA GAC CCC	336
Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro	
100 105 110	

CGG CGT AGG TCG CGT AAT TTG GGT  
 Arg Arg Arg Ser Arg Asn Leu Gly  
           115                          120

360

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 120 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn  
   1                          5                          10                          15  
 Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly  
                           20                          25                          30  
 Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala  
                           35                          40                          45  
 Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro  
   50                          55                          60  
*Q1. Cmp* Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Thr Trp Ala Gln Pro Gly  
   65                          70                          75                          80  
 Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Met Gly Trp Ala Gly Trp  
                           85                          90                          95  
 Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro  
                           100                          105                          110  
 Arg Arg Arg Ser Arg Asn Leu Gly  
   115                          120

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 450 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..450

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATG AGC ACA AAT CCT AAA CCT CAA AGA AAA ACC AAA CGT AAC ACC AAC	48
Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn	
1 5 10 15	
CGC CGC CCA CGG GAC GTT AAA TTC CCG GGC GGT GGT CAG ATC GTT GGT	96
Arg Arg Pro Arg Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly	
20 25 30	
GGA GTT TAC CTG TTG CCG CGC AGG GGC CCC AGG TTG GGT GTG CGC GCG	144
Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala	
35 40 45	
ACT AGG AAG ACT TCC GAG CGG TCG CAA CCT CGT GGA AGG CGA CAA CCT	192
Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro	
50 55 60	
ATC CCC AAG GCT CGC CGG CCC GAG GGT AGG ACC TGG GCT CAG CCC GGG	240
Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Thr Trp Ala Gln Pro Gly	
65 70 75 80	
TAC CCT TGG CCC CTC TAT GGC AAC GAG GGT ATG GGG TGG GCA GGA TGG	288
Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Met Gly Trp Ala Gly Trp	
85 90 95	
CTC CTG TCA CCC CGT GGC TCC CGG CCT AGT TGG GGC CCC ACG GAC CCC	336
Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro	
100 105 110	
CGG CGT AGG TCA CGC AAT TTG GGT AAG GTC ATC GAT ACC CTC ACA TGC	384
Arg Arg Arg Ser Arg Asn Leu Gly Lys Val Ile Asp Thr Leu Thr Cys	
115 120 125	
GGC TTC GCC GAC CTC ATG GGG TAC ATT CCG CTT GTC GGC GCC CCC CTA	432
Gly Phe Ala Asp Leu Met Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu	
130 135 140	
GGG GGC GCT GCC AGG GCC	450
Gly Gly Ala Ala Arg Ala	
145 150	

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 150 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn  
1 5 10 15  
Arg Arg Pro Arg Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly  
20 25 30  
Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala  
35 40 45  
Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro  
50 55 60  
Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Thr Trp Ala Gln Pro Gly  
65 70 75 80  
a) Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Met Gly Trp Ala Gly Trp  
c) 85 90 95  
Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro  
100 105 110  
Arg Arg Arg Ser Arg Asn Leu Gly Lys Val Ile Asp Thr Leu Thr Cys  
115 120 125  
Gly Phe Ala Asp Leu Met Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu  
130 135 140  
Gly Gly Ala Ala Arg Ala  
145 150

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 483 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
 (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 1..483

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

ATG GCT AGC GAA TTC ATG AGC ACA AAT CCT AAA CCT CAA AGA AAA ACC	48
Met Ala Ser Glu Phe Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr	
1 5 10 15	
AAA CGT AAC ACC AAC CGC CGC CCA CAG GAC GTT AAG TTC CCG GGC GGT	96
Lys Arg Asn Thr Asn Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly	
20 25 30	
GGT CAG ATC GTT GGT GGA GTT TAC CTG TTG CCG CGC AGG GGC CCC AGG	144
Gly Gln Ile Val Gly Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg	
35 40 45	
TTG GGT GTG CGC GCG ACT AGG AAG ACT TCC GAG CGG TCG CAA CCT CGT	192
Leu Gly Val Arg Ala Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg	
50 55 60	
GGA AGG CGA CAA CCT ATC CCC AAG GCT CGC CGG CCC GAG GGT AGG ACC	240
Gly Arg Arg Gln Pro Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Thr	
65 70 75 80	
TGG GCT CAG CCC GGG TAC CCT TGG CCC CTC TAT GGC AAC GAG GGT ATG	288
Trp Ala Gln Pro Gly Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Met	
85 90 95	
GGG TGG GCA GGA TGG CTC CTG TCA CCC CGT GGC TCT CGG CCT AGT TGG	336
Gly Trp Ala Gly Trp Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp	
100 105 110	
GGC CCC ACA GAC CCC CGG CGT AGG TCG CGT AAT TTG GGT GGA TCC AGA	384
Gly Pro Thr Asp Pro Arg Arg Arg Ser Arg Asn Leu Gly Gly Ser Arg	
115 120 125	
CGA CGA GGC AGG TCC CCT AGA AGA AGA ACT CCC TCG CCT CGC AGA CGA	432
Arg Arg Gly Arg Ser Pro Arg Arg Arg Thr Pro Ser Pro Arg Arg Arg	
130 135 140	
AGG TCT AAA TCG CCG CGT CGC AGA AGA TCT CAA TCT CGG GAA TCT CAA	480
Arg Ser Lys Ser Pro Arg Arg Arg Arg Ser Gln Ser Arg Glu Ser Gln	





(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 573 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..573

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

ATG GCT AGC GAA TTC ATG AGC ACA AAT CCT AAA CCT CAA AGA AAA ACC	48
Met Ala Ser Glu Phe Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr	
1 5 10 15	
AAA CGT AAC ACC AAC CGC CGC CCA CGG GAC GTT AAA TTC CCG GGC GGT	96
Lys Arg Asn Thr Asn Arg Arg Pro Arg Asp Val Lys Phe Pro Gly Gly	
20 25 30	
GGT CAG ATC GTT GGT GGA GTT TAC CTG TTG CCG CGC AGG GGC CCC AGG	144
Gly Gln Ile Val Gly Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg	
35 40 45	
TTG GGT GTG CGC GCG ACT AGG AAG ACT TCC GAG CGG TCG CAA CCT CGT	192
Leu Gly Val Arg Ala Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg	
50 55 60	
GGA AGG CGA CAA CCT ATC CCC AAG GCT CGC CGG CCC GAG GGT AGG ACC	240
Gly Arg Arg Gln Pro Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Thr	
65 70 75 80	
TGG GCT CAG CCC GGG TAC CCT TGG CCC CTC TAT GGC AAC GAG GGT ATG	288
Trp Ala Gln Pro Gly Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Met	
85 90 95	
GGG TGG GCA GGA TGG CTC CTG TCA CCC CGT GGC TCC CGG CCT AGT TGG	336
Gly Trp Ala Gly Trp Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp	
100 105 110	

GGC	CCC	ACG	GAC	CCC	CGG	CGT	AGG	TCA	CGC	AAT	TTG	GGT	AAG	GTC	ATC	384
Gly	Pro	Thr	Asp	Pro	Arg	Arg	Arg	Ser	Arg	Asn	Leu	Gly	Lys	Val	Ile	
		115					120					125				
GAT	ACC	CTC	ACA	TGC	GGC	TTC	GCC	GAC	CTC	ATG	GGG	TAC	ATT	CCG	CTT	432
Asp	Thr	Leu	Thr	Cys	Gly	Phe	Ala	Asp	Leu	Met	Gly	Tyr	Ile	Pro	Leu	
		130				135					140					
GTC	GGC	GCC	CCC	CTA	GGG	GGC	GCT	GCC	AGG	GCC	GGA	TCC	AGA	CGA	CGA	480
Val	Gly	Ala	Pro	Leu	Gly	Gly	Ala	Ala	Arg	Ala	Gly	Ser	Arg	Arg	Arg	
145					150					155					160	
GGC	AGG	TCC	CCT	AGA	AGA	AGA	ACT	CCC	TCG	CCT	CGC	AGA	CGA	AGG	TCT	528
Gly	Arg	Ser	Pro	Arg	Arg	Arg	Thr	Pro	Ser	Pro	Arg	Arg	Arg	Arg	Ser	
				165					170					175		
AAA	TCG	CCG	CGT	CGC	AGA	AGA	TCT	CAA	TCT	CGG	GAA	TCT	CAA	TGT		573
Lys	Ser	Pro	Arg	Arg	Arg	Arg	Ser	Gln	Ser	Arg	Glu	Ser	Gln	Cys		
			180					185					190			

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 191 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met	Ala	Ser	Glu	Phe	Met	Ser	Thr	Asn	Pro	Lys	Pro	Gln	Arg	Lys	Thr
1				5					10					15	
Lys	Arg	Asn	Thr	Asn	Arg	Arg	Pro	Arg	Asp	Val	Lys	Phe	Pro	Gly	Gly
			20					25					30		
Gly	Gln	Ile	Val	Gly	Gly	Val	Tyr	Leu	Leu	Pro	Arg	Arg	Gly	Pro	Arg
		35					40					45			
Leu	Gly	Val	Arg	Ala	Thr	Arg	Lys	Thr	Ser	Glu	Arg	Ser	Gln	Pro	Arg
	50					55					60				
Gly	Arg	Arg	Gln	Pro	Ile	Pro	Lys	Ala	Arg	Arg	Pro	Glu	Gly	Arg	Thr
65					70					75				80	
Trp	Ala	Gln	Pro	Gly	Tyr	Pro	Trp	Pro	Leu	Tyr	Gly	Asn	Glu	Gly	Met
				85					90					95	

Gly	Trp	Ala	Gly	Trp	Leu	Leu	Ser	Pro	Arg	Gly	Ser	Arg	Pro	Ser	Trp
			100					105					110		
Gly	Pro	Thr	Asp	Pro	Arg	Arg	Arg	Ser	Arg	Asn	Leu	Gly	Lys	Val	Ile
		115					120					125			
Asp	Thr	Leu	Thr	Cys	Gly	Phe	Ala	Asp	Leu	Met	Gly	Tyr	Ile	Pro	Leu
	130					135					140				
Val	Gly	Ala	Pro	Leu	Gly	Gly	Ala	Ala	Arg	Ala	Gly	Ser	Arg	Arg	Arg
145					150					155					160
Gly	Arg	Ser	Pro	Arg	Arg	Arg	Thr	Pro	Ser	Pro	Arg	Arg	Arg	Arg	Ser
				165					170					175	
Lys	Ser	Pro	Arg	Arg	Arg	Arg	Ser	Gln	Ser	Arg	Glu	Ser	Gln	Cys	
			180					185					190		

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 843 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..843

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ATG	GCT	AGC	ATG	AGC	ACA	AAT	CCT	AAA	CCT	CAA	AGA	AAA	ACC	AAA	CGT	48
Met	Ala	Ser	Met	Ser	Thr	Asn	Pro	Lys	Pro	Gln	Arg	Lys	Thr	Lys	Arg	
1				5				10						15		
AAC	ACC	AAC	CGC	CGC	CCA	CAG	GAC	GTT	AAG	TTC	CCG	GGC	GGT	GGT	CAG	96
Asn	Thr	Asn	Arg	Arg	Pro	Gln	Asp	Val	Lys	Phe	Pro	Gly	Gly	Gly	Gln	
			20					25					30			
ATC	GTT	GGT	GGA	GTT	TAC	CTG	TTG	CCG	CGC	AGG	GGC	CCC	AGG	TTG	GGT	144
Ile	Val	Gly	Gly	Val	Tyr	Leu	Leu	Pro	Arg	Arg	Gly	Pro	Arg	Leu	Gly	
		35					40					45				

GTG	CGC	GCG	ACT	AGG	AAG	ACT	TCC	GAG	CGG	TCG	CAA	CCT	CGT	GGA	AGG	192
Val	Arg	Ala	Thr	Arg	Lys	Thr	Ser	Glu	Arg	Ser	Gln	Pro	Arg	Gly	Arg	
	50					55					60					
CGA	CAA	CCT	ATC	CCC	AAG	GCT	CGC	CGG	CCC	GAG	GGT	AGG	ACC	TGG	GCT	240
Arg	Gln	Pro	Ile	Pro	Lys	Ala	Arg	Arg	Pro	Glu	Gly	Arg	Thr	Trp	Ala	
65					70					75					80	
CAG	CCC	GGG	TAC	CCT	TGG	CCC	CTC	TAT	GGC	AAC	GAG	GGT	ATG	GGG	TGG	288
Gln	Pro	Gly	Tyr	Pro	Trp	Pro	Leu	Tyr	Gly	Asn	Glu	Gly	Met	Gly	Trp	
				85					90					95		
GCA	GGA	TGG	CTC	CTG	TCA	CCC	CGT	GGC	TCT	CGG	CCT	AGT	TGG	GGC	CCC	336
Ala	Gly	Trp	Leu	Leu	Ser	Pro	Arg	Gly	Ser	Arg	Pro	Ser	Trp	Gly	Pro	
			100					105					110			
ACA	GAC	CCC	CGG	CGT	AGG	TCG	CGT	AAT	TTG	GGT	GAA	TTC	AGA	CGA	CGA	384
Thr	Asp	Pro	Arg	Arg	Arg	Ser	Arg	Asn	Leu	Gly	Glu	Phe	Arg	Arg	Arg	
		115					120					125				
GGC	AGG	TCC	CCT	AGA	AGA	AGA	ACT	CCC	TCG	CCT	CGC	AGA	CGA	AGG	TCT	432
Gly	Arg	Ser	Pro	Arg	Arg	Arg	Thr	Pro	Ser	Pro	Arg	Arg	Arg	Arg	Ser	
	130					135					140					
AAA	TCG	CCG	CGT	CGC	AGA	AGA	TCT	CAA	TCT	CGG	GAA	TCT	CAA	TGT	GAA	480
Lys	Ser	Pro	Arg	Arg	Arg	Arg	Ser	Gln	Ser	Arg	Glu	Ser	Gln	Cys	Glu	
145					150					155					160	
TTC	ATG	AGC	ACA	AAT	CCT	AAA	CCT	CAA	AGA	AAA	ACC	AAA	CGT	AAC	ACC	528
Phe	Met	Ser	Thr	Asn	Pro	Lys	Pro	Gln	Arg	Lys	Thr	Lys	Arg	Asn	Thr	
				165					170					175		
AAC	CGC	CGC	CCA	CAG	GAC	GTT	AAG	TTC	CCG	GGC	GGT	GGT	CAG	ATC	GTT	576
Asn	Arg	Arg	Pro	Gln	Asp	Val	Lys	Phe	Pro	Gly	Gly	Gly	Gln	Ile	Val	
			180					185					190			
GGT	GGA	GTT	TAC	CTG	TTG	CCG	CGC	AGG	GGC	CCC	AGG	TTG	GGT	GTG	CGC	624
Gly	Gly	Val	Tyr	Leu	Leu	Pro	Arg	Arg	Gly	Pro	Arg	Leu	Gly	Val	Arg	
		195					200					205				
GCG	ACT	AGG	AAG	ACT	TCC	GAG	CGG	TCG	CAA	CCT	CGT	GGA	AGG	CGA	CAA	672
Ala	Thr	Arg	Lys	Thr	Ser	Glu	Arg	Ser	Gln	Pro	Arg	Gly	Arg	Arg	Gln	
	210					215					220					
CCT	ATC	CCC	AAG	GCT	CGC	CGG	CCC	GAG	GGT	AGG	ACC	TGG	GCT	CAG	CCC	720
Pro	Ile	Pro	Lys	Ala	Arg	Arg	Pro	Glu	Gly	Arg	Thr	Trp	Ala	Gln	Pro	
225					230					235					240	
GGG	TAC	CCT	TGG	CCC	CTC	TAT	GGC	AAC	GAG	GGT	ATG	GGG	TGG	GCA	GGA	768

Gly	Tyr	Pro	Trp	Pro	Leu	Tyr	Gly	Asn	Glu	Gly	Met	Gly	Trp	Ala	Gly	
				245					250					255		
TGG	CTC	CTG	TCA	CCC	CGT	GGC	TCT	CGG	CCT	AGT	TGG	GGC	CCC	ACA	GAC	816
Trp	Leu	Leu	Ser	Pro	Arg	Gly	Ser	Arg	Pro	Ser	Trp	Gly	Pro	Thr	Asp	
			260					265					270			
CCC	CGG	CGT	AGG	TCG	CGT	AAT	TTG	GGT								843
Pro	Arg	Arg	Arg	Ser	Arg	Asn	Leu	Gly								
			275				280									

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 281 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Ala Ser Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg  
 1 5 10 15  
 Asn Thr Asn Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln  
 20 25 30  
 Ile Val Gly Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly  
 35 40 45  
 Val Arg Ala Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg  
 50 55 60  
 Arg Gln Pro Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Thr Trp Ala  
 65 70 75 80  
 Gln Pro Gly Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Met Gly Trp  
 85 90 95  
 Ala Gly Trp Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro  
 100 105 110  
 Thr Asp Pro Arg Arg Arg Ser Arg Asn Leu Gly Glu Phe Arg Arg Arg  
 115 120 125  
 Gly Arg Ser Pro Arg Arg Arg Thr Pro Ser Pro Arg Arg Arg Ser  
 130 135 140

Lys	Ser	Pro	Arg	Arg	Arg	Arg	Ser	Gln	Ser	Arg	Glu	Ser	Gln	Cys	Glu
145					150					155					160
Phe	Met	Ser	Thr	Asn	Pro	Lys	Pro	Gln	Arg	Lys	Thr	Lys	Arg	Asn	Thr
				165					170					175	
Asn	Arg	Arg	Pro	Gln	Asp	Val	Lys	Phe	Pro	Gly	Gly	Gly	Gln	Ile	Val
			180					185					190		
Gly	Gly	Val	Tyr	Leu	Leu	Pro	Arg	Arg	Gly	Pro	Arg	Leu	Gly	Val	Arg
		195					200					205			
Ala	Thr	Arg	Lys	Thr	Ser	Glu	Arg	Ser	Gln	Pro	Arg	Gly	Arg	Arg	Gln
	210					215					220				
Pro	Ile	Pro	Lys	Ala	Arg	Arg	Pro	Glu	Gly	Arg	Thr	Trp	Ala	Gln	Pro
225					230					235					240
Gly	Tyr	Pro	Trp	Pro	Leu	Tyr	Gly	Asn	Glu	Gly	Met	Gly	Trp	Ala	Gly
				245					250					255	
Trp	Leu	Leu	Ser	Pro	Arg	Gly	Ser	Arg	Pro	Ser	Trp	Gly	Pro	Thr	Asp
			260					265					270		
Pro	Arg	Arg	Arg	Ser	Arg	Asn	Leu	Gly							
			275				280								

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1245 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1245

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

ATG	GGC	TCG	TCT	CAT	CAT	GAG	ACG	CAC	TAT	GGC	TAT	GCG	ACG	CTA	AGC
Met	Gly	Ser	Ser	His	His	Glu	Thr	His	Tyr	Gly	Tyr	Ala	Thr	Leu	Ser
1				5					10					15	

TAT	GCG	GAC	TAC	TGG	GCC	GGG	GAG	TTG	GGG	CAG	AGT	AGG	GAC	GTG	CTT	96
Tyr	Ala	Asp	Tyr	Trp	Ala	Gly	Glu	Leu	Gly	Gln	Ser	Arg	Asp	Val	Leu	
			20					25					30			
TTG	GCG	GGT	AAT	GCC	GAG	GCG	GAC	CGC	GCG	GGG	GAT	CTC	GAC	GCA	GGC	144
Leu	Ala	Gly	Asn	Ala	Glu	Ala	Asp	Arg	Ala	Gly	Asp	Leu	Asp	Ala	Gly	
		35					40					45				
ATG	TTC	GAT	GCA	GTT	TCT	CGC	GCA	ACC	CAC	GGG	CAT	GGC	GCG	TTC	CGT	192
Met	Phe	Asp	Ala	Val	Ser	Arg	Ala	Thr	His	Gly	His	Gly	Ala	Phe	Arg	
	50					55					60					
CAG	CAA	TTT	CAG	TAC	GCG	GTT	GAG	GTA	TTG	GGC	GAA	AAG	GTT	CTC	TCG	240
Gln	Gln	Phe	Gln	Tyr	Ala	Val	Glu	Val	Leu	Gly	Glu	Lys	Val	Leu	Ser	
65					70					75					80	
AAG	CAG	GAG	ACC	GAA	GAC	AGC	AGG	GGA	AGA	AAA	AAG	TGG	GAG	TAC	GAG	288
Lys	Gln	Glu	Thr	Glu	Asp	Ser	Arg	Gly	Arg	Lys	Lys	Trp	Glu	Tyr	Glu	
				85				90						95		
ACT	GAC	CCA	AGC	GTT	ACT	AAG	ATG	GTG	CGT	GCC	TCT	GCG	TCA	TTT	CAG	336
Thr	Asp	Pro	Ser	Val	Thr	Lys	Met	Val	Arg	Ala	Ser	Ala	Ser	Phe	Gln	
			100					105					110			
GAT	TTG	GGA	GAG	GAC	GGG	GAG	ATT	AAG	TTT	GAA	GCA	GTC	GAG	GGT	GCA	384
Asp	Leu	Gly	Glu	Asp	Gly	Glu	Ile	Lys	Phe	Glu	Ala	Val	Glu	Gly	Ala	
		115					120					125				
GTG	GCG	TTG	GCG	GAT	CGC	GCG	AGT	TCC	TTC	ATG	GTT	GAC	AGC	GAG	GAA	432
Val	Ala	Leu	Ala	Asp	Arg	Ala	Ser	Ser	Phe	Met	Val	Asp	Ser	Glu	Glu	
	130					135					140					
TAC	AAG	ATT	ACG	AAC	GTA	AAG	GTT	CAC	GGT	ATG	AAG	TTT	GTC	CCA	GTT	480
Tyr	Lys	Ile	Thr	Asn	Val	Lys	Val	His	Gly	Met	Lys	Phe	Val	Pro	Val	
145					150					155					160	
GCG	GTT	CCT	CAT	GAA	TTA	AAA	GGG	ATT	GCA	AAG	GAG	AAG	TTT	CAC	TTC	528
Ala	Val	Pro	His	Glu	Leu	Lys	Gly	Ile	Ala	Lys	Glu	Lys	Phe	His	Phe	
				165				170					175			
GTG	GAA	GAC	TCC	CGC	GTT	ACG	GAG	AAT	ACC	AAC	GGC	CTT	AAG	ACA	ATG	576
Val	Glu	Asp	Ser	Arg	Val	Thr	Glu	Asn	Thr	Asn	Gly	Leu	Lys	Thr	Met	
			180					185					190			
CTC	ACT	GAG	GAT	AGT	TTT	TCT	GCA	CGT	AAG	GTA	AGC	AGC	ATG	GAG	AGC	624
Leu	Thr	Glu	Asp	Ser	Phe	Ser	Ala	Arg	Lys	Val	Ser	Ser	Met	Glu	Ser	
		195					200					205				
CCG	CAC	GAC	CTT	GTG	GTA	GAC	ACG	GTG	GGT	ACC	GTC	TAC	CAC	AGC	CGT	672

Pro	His	Asp	Leu	Val	Val	Asp	Thr	Val	Gly	Thr	Val	Tyr	His	Ser	Arg	
210						215					220					
TTT	GGT	TCG	GAC	GCA	GAG	GCT	TCT	GTG	ATG	CTG	AAA	AGG	GCT	GAT	GGC	720
Phe	Gly	Ser	Asp	Ala	Glu	Ala	Ser	Val	Met	Leu	Lys	Arg	Ala	Asp	Gly	
225					230					235					240	
TCT	GAG	CTG	TCG	CAC	CGT	GAG	TTC	ATC	GAC	TAT	GTG	ATG	AAC	TTC	AAC	768
Ser	Glu	Leu	Ser	His	Arg	Glu	Phe	Ile	Asp	Tyr	Val	Met	Asn	Phe	Asn	
				245					250					255		
ACG	GTC	CGC	TAC	GAC	TAC	TAC	GGT	GAT	GAC	GCG	AGC	TAC	ACC	AAT	CTG	816
Thr	Val	Arg	Tyr	Asp	Tyr	Tyr	Gly	Asp	Asp	Ala	Ser	Tyr	Thr	Asn	Leu	
				260				265					270			
ATG	GCG	AGT	TAT	GGC	ACC	AAG	CAC	TCT	GCT	GAC	TCC	TGG	TGG	AAG	ACA	864
Met	Ala	Ser	Tyr	Gly	Thr	Lys	His	Ser	Ala	Asp	Ser	Trp	Trp	Lys	Thr	
		275					280					285				
GGA	AGA	GTG	CCC	CGC	ATT	TCG	TGT	GGT	ATC	AAC	TAT	GGG	TTC	GAT	CGG	912
Gly	Arg	Val	Pro	Arg	Ile	Ser	Cys	Gly	Ile	Asn	Tyr	Gly	Phe	Asp	Arg	
	290					295					300					
TTT	AAA	GGT	TCA	GGG	CCG	GGA	TAC	TAC	AGG	CTG	ACT	TTG	ATT	GCG	AAC	960
Phe	Lys	Gly	Ser	Gly	Pro	Gly	Tyr	Tyr	Arg	Leu	Thr	Leu	Ile	Ala	Asn	
305					310					315					320	
GGG	TAT	AGG	GAC	GTA	GTT	GCT	GAT	GTG	CGC	TTC	CTT	CCC	AAG	TAC	GAG	1008
Gly	Tyr	Arg	Asp	Val	Val	Ala	Asp	Val	Arg	Phe	Leu	Pro	Lys	Tyr	Glu	
				325					330					335		
GGG	AAC	ATC	GAT	ATT	GGG	TTG	AAG	GGG	AAG	GTG	CTG	ACC	ATA	GGG	GGC	1056
Gly	Asn	Ile	Asp	Ile	Gly	Leu	Lys	Gly	Lys	Val	Leu	Thr	Ile	Gly	Gly	
			340					345					350			
GCG	GAC	GCG	GAG	ACT	CTG	ATG	GAT	GCT	GCA	GTT	GAC	GTG	TTT	GCC	GAT	1104
Ala	Asp	Ala	Glu	Thr	Leu	Met	Asp	Ala	Ala	Val	Asp	Val	Phe	Ala	Asp	
		355					360					365				
GGA	CAG	CCT	AAG	CTT	GTC	AGC	GAT	CAA	GCG	GTG	AGC	TTG	GGG	CAG	AAT	1152
Gly	Gln	Pro	Lys	Leu	Val	Ser	Asp	Gln	Ala	Val	Ser	Leu	Gly	Gln	Asn	
	370					375					380					
GTC	CTC	TCT	GCG	GAT	TTC	ACT	CCC	GGC	ACT	GAG	TAC	ACG	GTT	GAG	GTT	1200
Val	Leu	Ser	Ala	Asp	Phe	Thr	Pro	Gly	Thr	Glu	Tyr	Thr	Val	Glu	Val	
385					390					395					400	
AGG	TTC	AAG	GAA	TTT	GGT	TCT	GTG	CGT	GCG	AAG	GTA	GTG	GCC	CAG		1245
Arg	Phe	Lys	Glu	Phe	Gly	Ser	Val	Arg	Ala	Lys	Val	Val	Ala	Gln		



405

410

415

## (2) INFORMATION FOR SEQ ID NO:14:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 415 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Gly Ser Ser His His Glu Thr His Tyr Gly Tyr Ala Thr Leu Ser  
 1 5 10 15  
 Tyr Ala Asp Tyr Trp Ala Gly Glu Leu Gly Gln Ser Arg Asp Val Leu  
 20 25 30  
 Leu Ala Gly Asn Ala Glu Ala Asp Arg Ala Gly Asp Leu Asp Ala Gly  
 35 40 45  
 Met Phe Asp Ala Val Ser Arg Ala Thr His Gly His Gly Ala Phe Arg  
 50 55 60  
 Gln Gln Phe Gln Tyr Ala Val Glu Val Leu Gly Glu Lys Val Leu Ser  
 65 70 75 80  
 Lys Gln Glu Thr Glu Asp Ser Arg Gly Arg Lys Lys Trp Glu Tyr Glu  
 85 90 95  
 Thr Asp Pro Ser Val Thr Lys Met Val Arg Ala Ser Ala Ser Phe Gln  
 100 105 110  
 Asp Leu Gly Glu Asp Gly Glu Ile Lys Phe Glu Ala Val Glu Gly Ala  
 115 120 125  
 Val Ala Leu Ala Asp Arg Ala Ser Ser Phe Met Val Asp Ser Glu Glu  
 130 135 140  
 Tyr Lys Ile Thr Asn Val Lys Val His Gly Met Lys Phe Val Pro Val  
 145 150 155 160  
 Ala Val Pro His Glu Leu Lys Gly Ile Ala Lys Glu Lys Phe His Phe  
 165 170 175  
 Val Glu Asp Ser Arg Val Thr Glu Asn Thr Asn Gly Leu Lys Thr Met  
 180 185 190

Leu	Thr	Glu	Asp	Ser	Phe	Ser	Ala	Arg	Lys	Val	Ser	Ser	Met	Glu	Ser			
		195					200					205						
Pro	His	Asp	Leu	Val	Val	Asp	Thr	Val	Gly	Thr	Val	Tyr	His	Ser	Arg			
	210					215					220							
Phe	Gly	Ser	Asp	Ala	Glu	Ala	Ser	Val	Met	Leu	Lys	Arg	Ala	Asp	Gly			
225					230					235					240			
Ser	Glu	Leu	Ser	His	Arg	Glu	Phe	Ile	Asp	Tyr	Val	Met	Asn	Phe	Asn			
				245					250					255				
Thr	Val	Arg	Tyr	Asp	Tyr	Tyr	Gly	Asp	Asp	Ala	Ser	Tyr	Thr	Asn	Leu			
			260					265					270					
Met	Ala	Ser	Tyr	Gly	Thr	Lys	His	Ser	Ala	Asp	Ser	Trp	Trp	Lys	Thr			
		275					280					285						
Gly	Arg	Val	Pro	Arg	Ile	Ser	Cys	Gly	Ile	Asn	Tyr	Gly	Phe	Asp	Arg			
	290					295					300							
Phe	Lys	Gly	Ser	Gly	Pro	Gly	Tyr	Tyr	Arg	Leu	Thr	Leu	Ile	Ala	Asn			
305					310					315					320			
Gly	Tyr	Arg	Asp	Val	Val	Ala	Asp	Val	Arg	Phe	Leu	Pro	Lys	Tyr	Glu			
				325					330					335				
Gly	Asn	Ile	Asp	Ile	Gly	Leu	Lys	Gly	Lys	Val	Leu	Thr	Ile	Gly	Gly			
			340					345					350					
Ala	Asp	Ala	Glu	Thr	Leu	Met	Asp	Ala	Ala	Val	Asp	Val	Phe	Ala	Asp			
		355					360					365						
Gly	Gln	Pro	Lys	Leu	Val	Ser	Asp	Gln	Ala	Val	Ser	Leu	Gly	Gln	Asn			
	370					375					380							
Val	Leu	Ser	Ala	Asp	Phe	Thr	Pro	Gly	Thr	Glu	Tyr	Thr	Val	Glu	Val			
385					390					395					400			
Arg	Phe	Lys	Glu	Phe	Gly	Ser	Val	Arg	Ala	Lys	Val	Val	Ala	Gln				
			405						410					415				

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1368 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
 (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 1..1368

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

ATG GCT AGC GAA TTC ATG GGC TCG TCT CAT CAT GAG ACG CAC TAT GGC	48
Met Ala Ser Glu Phe Met Gly Ser Ser His His Glu Thr His Tyr Gly	
1 5 10 15	
TAT GCG ACG CTA AGC TAT GCG GAC TAC TGG GCC GGG GAG TTG GGG CAG	96
Tyr Ala Thr Leu Ser Tyr Ala Asp Tyr Trp Ala Gly Glu Leu Gly Gln	
20 25 30	
AGT AGG GAC GTG CTT TTG GCG GGT AAT GCC GAG GCG GAC CGC GCG GGG	144
Ser Arg Asp Val Leu Leu Ala Gly Asn Ala Glu Ala Asp Arg Ala Gly	
35 40 45	
GAT CTC GAC GCA GGC ATG TTC GAT GCA GTT TCT CGC GCA ACC CAC GGG	192
Asp Leu Asp Ala Gly Met Phe Asp Ala Val Ser Arg Ala Thr His Gly	
50 55 60	
CAT GGC GCG TTC CGT CAG CAA TTT CAG TAC GCG GTT GAG GTA TTG GGC	240
His Gly Ala Phe Arg Gln Gln Phe Gln Tyr Ala Val Glu Val Leu Gly	
65 70 75 80	
GAA AAG GTT CTC TCG AAG CAG GAG ACC GAA GAC AGC AGG GGA AGA AAA	288
Glu Lys Val Leu Ser Lys Gln Glu Thr Glu Asp Ser Arg Gly Arg Lys	
85 90 95	
AAG TGG GAG TAC GAG ACT GAC CCA AGC GTT ACT AAG ATG GTG CGT GCC	336
Lys Trp Glu Tyr Glu Thr Asp Pro Ser Val Thr Lys Met Val Arg Ala	
100 105 110	
TCT GCG TCA TTT CAG GAT TTG GGA GAG GAC GGG GAG ATT AAG TTT GAA	384
Ser Ala Ser Phe Gln Asp Leu Gly Glu Asp Gly Glu Ile Lys Phe Glu	
115 120 125	
GCA GTC GAG GGT GCA GTA GCG TTG GCG GAT CGC GCG AGT TCC TTC ATG	432
Ala Val Glu Gly Ala Val Ala Leu Ala Asp Arg Ala Ser Ser Phe Met	
130 135 140	
GTT GAC AGC GAG GAA TAC AAG ATT ACG AAC GTA AAG GTT CAC GGT ATG	480
Val Asp Ser Glu Glu Tyr Lys Ile Thr Asn Val Lys Val His Gly Met	

145						150						155						160	
AAG TTT GTC CCA GTT GCG GTT CCT CAT GAA TTA AAA GGG ATT GCA AAG																			528
Lys Phe Val Pro Val Ala Val Pro His Glu Leu Lys Gly Ile Ala Lys						165						170						175	
GAG AAG TTT CAC TTC GTG GAA GAC TCC CGC GTT ACG GAG AAT ACC AAC																			576
Glu Lys Phe His Phe Val Glu Asp Ser Arg Val Thr Glu Asn Thr Asn						180						185						190	
GGC CTT AAG ACA ATG CTC ACT GAG GAT AGT TTT TCT GCA CGT AAG GTA																			624
Gly Leu Lys Thr Met Leu Thr Glu Asp Ser Phe Ser Ala Arg Lys Val						195						200						205	
AGC AGC ATG GAG AGC CCG CAC GAC CTT GTG GTA GAC ACG GTG GGT ACC																			672
Ser Ser Met Glu Ser Pro His Asp Leu Val Val Asp Thr Val Gly Thr						210						215						220	
GTC TAC CAC AGC CGT TTT GGT TCG GAC GCA GAG GCT TCT GTG ATG CTG																			720
Val Tyr His Ser Arg Phe Gly Ser Asp Ala Glu Ala Ser Val Met Leu						225						230						235	240
AAA AGG GCT GAT GGC TCT GAG CTG TCG CAC CGT GAG TTC ATC GAC TAT																			768
Lys Arg Ala Asp Gly Ser Glu Leu Ser His Arg Glu Phe Ile Asp Tyr						245						250						255	
GTG ATG AAC TTC AAC ACG GTC CGC TAC GAC TAC TAC GGT GAT GAC GCG																			816
Val Met Asn Phe Asn Thr Val Arg Tyr Asp Tyr Tyr Gly Asp Asp Ala						260						265						270	
AGC TAC ACC AAT CTG ATG GCG AGT TAT GGC ACC AAG CAC TCT GCT GAC																			864
Ser Tyr Thr Asn Leu Met Ala Ser Tyr Gly Thr Lys His Ser Ala Asp						275						280						285	
TCC TGG TGG AAG ACA GGA AGA GTG CCC CGC ATT TCG TGT GGT ATC AAC																			912
Ser Trp Trp Lys Thr Gly Arg Val Pro Arg Ile Ser Cys Gly Ile Asn						290						295						300	
TAT GGG TTC GAT CGG TTT AAA GGT TCA GGG CCG GGA TAC TAC AGG CTG																			960
Tyr Gly Phe Asp Arg Phe Lys Gly Ser Gly Pro Gly Tyr Tyr Arg Leu						305						310						315	320
ACT TTG ATT GCG AAC GGG TAT AGG GAC GTA GTT GCT GAT GTG CGC TTC																			1008
Thr Leu Ile Ala Asn Gly Tyr Arg Asp Val Val Ala Asp Val Arg Phe						325						330						335	
CTT CCC AAG TAC GAG GGG AAC ATC GAT ATT GGG TTG AAG GGG AAG GTG																			1056
Leu Pro Lys Tyr Glu Gly Asn Ile Asp Ile Gly Leu Lys Gly Lys Val						340						345						350	

CTG	ACC	ATA	GGG	GGC	GCG	GAC	GCG	GAG	ACT	CTG	ATG	GAT	GCT	GCA	GTT	1104
Leu	Thr	Ile	Gly	Gly	Ala	Asp	Ala	Glu	Thr	Leu	Met	Asp	Ala	Ala	Val	
		355					360					365				
GAC	GTG	TTT	GCC	GAT	GGA	CAG	CCT	AAG	CTT	GTC	AGC	GAT	CAA	GCG	GTG	1152
Asp	Val	Phe	Ala	Asp	Gly	Gln	Pro	Lys	Leu	Val	Ser	Asp	Gln	Ala	Val	
	370					375					380					
AGC	TTG	GGG	CAG	AAT	GTC	CTC	TCT	GCG	GAT	TTC	ACT	CCC	GGC	ACT	GAG	1200
Ser	Leu	Gly	Gln	Asn	Val	Leu	Ser	Ala	Asp	Phe	Thr	Pro	Gly	Thr	Glu	
385					390					395					400	
TAC	ACG	GTT	GAG	GTT	AGG	TTC	AAG	GAA	TTT	GGT	TCT	GTG	CGT	GCG	AAG	1248
Tyr	Thr	Val	Glu	Val	Arg	Phe	Lys	Glu	Phe	Gly	Ser	Val	Arg	Ala	Lys	
			405						410					415		
GTA	GTG	GCC	CAG	GGA	TCC	AGA	CGA	CGA	GGC	AGG	TCC	CCT	AGA	AGA	AGA	1296
Val	Val	Ala	Gln	Gly	Ser	Arg	Arg	Arg	Gly	Arg	Ser	Pro	Arg	Arg	Arg	
			420					425					430			
ACT	CCC	TCG	CCT	CGC	AGA	CGA	AGG	TCT	AAA	TCG	CCG	CGT	CGC	AGA	AGA	1344
Thr	Pro	Ser	Pro	Arg	Arg	Arg	Arg	Ser	Lys	Ser	Pro	Arg	Arg	Arg	Arg	
		435					440					445				
TCT	CAA	TCT	CGG	GAA	TCT	CAA	TGT									1368
a! Ser	Gln	Ser	Arg	Glu	Ser	Gln	Cys									
6mV.	450					455										

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 456 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met	Ala	Ser	Glu	Phe	Met	Gly	Ser	Ser	His	His	Glu	Thr	His	Tyr	Gly
1				5					10					15	
Tyr	Ala	Thr	Leu	Ser	Tyr	Ala	Asp	Tyr	Trp	Ala	Gly	Glu	Leu	Gly	Gln
			20					25					30		
Ser	Arg	Asp	Val	Leu	Leu	Ala	Gly	Asn	Ala	Glu	Ala	Asp	Arg	Ala	Gly
		35					40					45			

Asp Leu Asp Ala Gly Met Phe Asp Ala Val Ser Arg Ala Thr His Gly  
 50 55 60  
 His Gly Ala Phe Arg Gln Gln Phe Gln Tyr Ala Val Glu Val Leu Gly  
 65 70 75 80  
 Glu Lys Val Leu Ser Lys Gln Glu Thr Glu Asp Ser Arg Gly Arg Lys  
 85 90 95  
 Lys Trp Glu Tyr Glu Thr Asp Pro Ser Val Thr Lys Met Val Arg Ala  
 100 105 110  
 Ser Ala Ser Phe Gln Asp Leu Gly Glu Asp Gly Glu Ile Lys Phe Glu  
 115 120 125  
 Ala Val Glu Gly Ala Val Ala Leu Ala Asp Arg Ala Ser Ser Phe Met  
 130 135 140  
 Val Asp Ser Glu Glu Tyr Lys Ile Thr Asn Val Lys Val His Gly Met  
 145 150 155 160  
 Lys Phe Val Pro Val Ala Val Pro His Glu Leu Lys Gly Ile Ala Lys  
 165 170 175  
 Glu Lys Phe His Phe Val Glu Asp Ser Arg Val Thr Glu Asn Thr Asn  
 180 185 190  
 Gly Leu Lys Thr Met Leu Thr Glu Asp Ser Phe Ser Ala Arg Lys Val  
 195 200 205  
 Ser Ser Met Glu Ser Pro His Asp Leu Val Val Asp Thr Val Gly Thr  
 210 215 220  
 Val Tyr His Ser Arg Phe Gly Ser Asp Ala Glu Ala Ser Val Met Leu  
 225 230 235 240  
 Lys Arg Ala Asp Gly Ser Glu Leu Ser His Arg Glu Phe Ile Asp Tyr  
 245 250 255  
 Val Met Asn Phe Asn Thr Val Arg Tyr Asp Tyr Tyr Gly Asp Asp Ala  
 260 265 270  
 Ser Tyr Thr Asn Leu Met Ala Ser Tyr Gly Thr Lys His Ser Ala Asp  
 275 280 285  
 Ser Trp Trp Lys Thr Gly Arg Val Pro Arg Ile Ser Cys Gly Ile Asn  
 290 295 300  
 Tyr Gly Phe Asp Arg Phe Lys Gly Ser Gly Pro Gly Tyr Tyr Arg Leu

305		310		315		320									
Thr	Leu	Ile	Ala	Asn	Gly	Tyr	Arg	Asp	Val	Val	Ala	Asp	Val	Arg	Phe
				325					330					335	
Leu	Pro	Lys	Tyr	Glu	Gly	Asn	Ile	Asp	Ile	Gly	Leu	Lys	Gly	Lys	Val
			340					345					350		
Leu	Thr	Ile	Gly	Gly	Ala	Asp	Ala	Glu	Thr	Leu	Met	Asp	Ala	Ala	Val
		355					360					365			
Asp	Val	Phe	Ala	Asp	Gly	Gln	Pro	Lys	Leu	Val	Ser	Asp	Gln	Ala	Val
	370					375					380				
Ser	Leu	Gly	Gln	Asn	Val	Leu	Ser	Ala	Asp	Phe	Thr	Pro	Gly	Thr	Glu
385					390					395					400
Tyr	Thr	Val	Glu	Val	Arg	Phe	Lys	Glu	Phe	Gly	Ser	Val	Arg	Ala	Lys
			405						410					415	
Val	Val	Ala	Gln	Gly	Ser	Arg	Arg	Arg	Gly	Arg	Ser	Pro	Arg	Arg	Arg
			420					425					430		
Thr	Pro	Ser	Pro	Arg	Arg	Arg	Arg	Ser	Lys	Ser	Pro	Arg	Arg	Arg	Arg
		435					440					445			
Ser	Gln	Ser	Arg	Glu	Ser	Gln	Cys								
	450					455									

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 153 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..153

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

ATG GCC AGA TAC CGA TGC TGC CGC AGC AAA AGC AGG AGC AGA TGC CGC

48

Met	Ala	Arg	Tyr	Arg	Cys	Cys	Arg	Ser	Lys	Ser	Arg	Ser	Arg	Cys	Arg		
1				5					10					15			
CGT	CGC	AGA	CGA	AGA	TGT	CGC	AGA	CGG	AGG	AGG	CGA	TGC	TGC	CGG	CGG		96
Arg	Arg	Arg	Arg	Arg	Cys	Arg	Arg	Arg	Arg	Arg	Arg	Cys	Cys	Arg	Arg		
			20					25					30				
AGG	AGG	CGA	AGA	TGC	TGC	CGT	CGC	CGC	CGC	TCA	TAC	ACC	ATA	AGG	TGT		144
Arg	Arg	Arg	Arg	Cys	Cys	Arg	Arg	Arg	Arg	Ser	Tyr	Thr	Ile	Arg	Cys		
			35					40					45				
AAA	AAA	TAC															153
Lys	Lys	Tyr															
			50														

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 51 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met	Ala	Arg	Tyr	Arg	Cys	Cys	Arg	Ser	Lys	Ser	Arg	Ser	Arg	Cys	Arg		
1				5					10					15			
Arg	Arg	Arg	Arg	Arg	Cys	Arg	Arg	Arg	Arg	Arg	Arg	Cys	Cys	Arg	Arg		
			20					25					30				
Arg	Arg	Arg	Arg	Cys	Cys	Arg	Arg	Arg	Arg	Ser	Tyr	Thr	Ile	Arg	Cys		
			35					40					45				
Lys	Lys	Tyr															
			50														

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 528 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"



(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..528

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

ATG GCT AGC ATG AGC ACA AAT CCT AAA CCT CAA AGA AAA ACC AAA CGT	48
Met Ala Ser Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg	
1 5 10 15	
AAC ACC AAC CGC CGC CCA CGG GAC GTT AAA TTC CCG GGC GGT GGT CAG	96
Asn Thr Asn Arg Arg Pro Arg Asp Val Lys Phe Pro Gly Gly Gly Gln	
20 25 30	
ATC GTT GGT GGA GTT TAC CTG TTG CCG CGC AGG GGC CCC AGG TTG GGT	144
Ile Val Gly Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly	
35 40 45	
GTG CGC GCG ACT AGG AAG ACT TCC GAG CGG TCG CAA CCT CGT GGA AGG	192
Val Arg Ala Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg	
50 55 60	
CGA CAA CCT ATC CCC AAG GCT CGC CGG CCC GAG GGT AGG ACC TGG GCT	240
Arg Gln Pro Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Thr Trp Ala	
65 70 75 80	
<i>a'</i> CAG CCC GGG TAC CCT TGG CCC CTC TAT GGC AAC GAG GGT ATG GGG TGG	288
<i>met</i> Gln Pro Gly Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Met Gly Trp	
85 90 95	
GCA GGA TGG CTC CTG TCA CCC CGT GGC TCC CGG CCT AGT TGG GGC CCC	336
Ala Gly Trp Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro	
100 105 110	
ACG GAC CCC CGG CGT AGG TCA CGC AAT TTG GGT GAA TTC ATG GCC AGA	384
Thr Asp Pro Arg Arg Arg Ser Arg Asn Leu Gly Glu Phe Met Ala Arg	
115 120 125	
TAC CGA TGC TGC CGC AGC AAA AGC AGG AGC AGA TGC CGC CGT CGC AGA	432
Tyr Arg Cys Cys Arg Ser Lys Ser Arg Ser Arg Cys Arg Arg Arg Arg	
130 135 140	
CGA AGA TGT CGC AGA CGG AGG AGG CGA TGC TGC CGG CGG AGG AGG CGA	480
Arg Arg Cys Arg Arg Arg Arg Arg Arg Cys Cys Arg Arg Arg Arg Arg	
145 150 155 160	
AGA TGC TGC CGT CGC CGC CGC TCA TAC ACC ATA AGG TGT AAA AAA TAC	528

Arg Cys Cys Arg Arg Arg Arg Ser Tyr Thr Ile Arg Cys Lys Lys Tyr  
165 170 175

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 176 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Met Ala Ser Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg  
1 5 10 15

Asn Thr Asn Arg Arg Pro Arg Asp Val Lys Phe Pro Gly Gly Gly Gln  
20 25 30

Ile Val Gly Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly  
35 40 45

Val Arg Ala Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg  
50 55 60

Arg Gln Pro Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Thr Trp Ala  
65 70 75 80

Gln Pro Gly Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Met Gly Trp  
85 90 95

Ala Gly Trp Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro  
100 105 110

Thr Asp Pro Arg Arg Arg Ser Arg Asn Leu Gly Glu Phe Met Ala Arg  
115 120 125

Tyr Arg Cys Cys Arg Ser Lys Ser Arg Ser Arg Cys Arg Arg Arg Arg  
130 135 140

Arg Arg Cys Arg Arg Arg Arg Arg Arg Cys Cys Arg Arg Arg Arg Arg  
145 150 155 160

Arg Cys Cys Arg Arg Arg Arg Ser Tyr Thr Ile Arg Cys Lys Lys Tyr  
165 170 175